Smooth maps from clumpy data

Marco Lombardi and Peter Schneider

Institüt für Astrophysik und Extraterrestrische Forschung, Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany

Abstract. We study an estimator for smoothing irregularly sampled data into a smooth map. The estimator has been widely used in astronomy, owing to its low level of noise; it involves a weight function – or smoothing kernel – $w(\boldsymbol{\theta})$. We show that this estimator is not unbiased, in the sense that the expectation value of the smoothed map is not the underlying process convolved with w, but a convolution with a modified kernel $w_{\mathrm{eff}}(\boldsymbol{\theta})$. We show how to calculate w_{eff} for a given kernel w and investigate its properties. In particular, it is found that (1) w_{eff} is normalized, (2) has a shape 'similar' to the original kernel w, (3) converges to w in the limit of high number density of data points, and (4) reduces to a top-hat filter in the limit of very small number density of data points. Hence, although the estimator is biased, the bias is well understood analytically, and since w_{eff} has all the desired properties of a smoothing kernel, the estimator is in fact very useful. We present explicit examples for several filter functions which are commonly used, and provide a series expression valid in the limit of large density of data points.

1. Introduction

A common problem in astronomy is the smoothing of some irregularly sampled data into a continuous map. It is hard to list all possible cases where such a problem is encountered. Just to give a few examples, we mention the determination of the stellar distribution of our Galaxy, the mapping of column density of dark molecular clouds from the absorption of background stars, the determination of the distribution function in globular clusters from radial and, recently, proper motions of stars, the determination of cosmological large-scale structures from redshift surveys, and the mass reconstruction in galaxy clusters from the observed distortion of background galaxies using weak lensing techniques. Similarly, one-dimensional data, such as time series of some events, often need to be smoothed in order to obtain a real function.

The use of a smooth map is convenient for at least two reasons. First, a smooth map can be better analyzed than irregularly sampled data. Second, if the smoothing is done in a sufficiently coarse way, the smooth map is significantly less noisy than the individual data. The drawback related to this last point is a loss in resolution, but often this is a price that we have to pay in order to obtain results with a decent signal-to-noise ratio.

In many cases the transformation of irregularly sampled data into a smooth map follows a standard approach. A positive weight function, describing the relative weight of a datum at the position $\boldsymbol{\theta} + \boldsymbol{\phi}$ on the point $\boldsymbol{\theta}$, is introduced. This function is generally of the form $w(\boldsymbol{\phi})$, i.e. is independent of absolute position $\boldsymbol{\theta}$ of the point, and actually often depends only on the separation $|\boldsymbol{\phi}|$. The weight function $w(\boldsymbol{\phi})$ is also chosen so that its value is large when the datum is close to the point, i.e. when $|\boldsymbol{\phi}|$ is small, and vanishes when $|\boldsymbol{\phi}|$ is large. Then, the data are averaged using a weighted mean with the weights given by the function w. More precisely, calling $\hat{f}_n$ the n-th datum obtained at the position $\boldsymbol{\theta}_n$, the smooth map is defined as

$$\tilde{f}(\boldsymbol{\theta}) = \frac{\sum_n \hat{f}_n w(\boldsymbol{\theta} - \boldsymbol{\theta}_n)}{\sum_n w(\boldsymbol{\theta} - \boldsymbol{\theta}_n)} \ . \tag{1}$$

It is *reasonable* to *assume* that this standard approach works well and produces good results, and the frequent use of this estimator suggests this is the case. On the other hand, in our opinion the properties of the smoothing should be better characterized by means of rigorous calculations. Some authors have actually already studied the smoothing using some approximations, and have obtained some preliminary results (e.g., Lombardi & Bertin 1998, van Waerbeke 2000). To our knowledge, however, the general problem has not been fully addressed so far and in particular there are no exact results known.

In this paper we consider in detail the effect of the smoothing on irregularly sampled data and derive a number of *exact* properties for the resulting map. We assume that the measurements $\hat{f}_n$ are *unbiased* estimates of some unknown field $f(\boldsymbol{\theta})$ at the positions $\boldsymbol{\theta}_n$, and we study the expectation value of the map $\tilde{f}(\boldsymbol{\theta})$ using an *ensemble average*, i.e. taking the N positions $\{\boldsymbol{\theta}_n\}$ as random variables. We then show that the expectation value for the smooth map of Eq. (1) is given by

$$\langle \tilde{f}(\boldsymbol{\theta}) \rangle = \int f(\boldsymbol{\theta}') w_{\mathrm{eff}}(\boldsymbol{\theta} - \boldsymbol{\theta}') \, \mathrm{d}^2\theta' \ . \tag{2}$$

Thus, $\langle \tilde{f} \rangle$ is the convolution of the unknown field f with an *effective weight* w_{eff} which, in general, differs from the weight function w. We also show that w_{eff} has a "similar" shape as w and converges to w for a large number of objects N, but in general w_{eff} is broader than w. Moreover, the effective weight is normalized, so that no signal is "lost" or "created." Finally, we obtain some analytical expansions for w_{eff}, and investigate its behavior in a number of interesting cases.

A common alternative to Eq. (1) is a non-normalized weighted sum, defined as

$$\tilde{f}(\boldsymbol{\theta}) = \frac{1}{\rho} \sum_{n=1}^{N} \hat{f}_n w(\boldsymbol{\theta} - \boldsymbol{\theta}_n) , \tag{3}$$

where w has been assumed to have unit integral [see below Eq. (6)]. The statistical properties of this estimator can be easily derived, and in particular we obtain for its expectation value [cf. Eq. (3)]

$$\langle \tilde{f}(\boldsymbol{\theta}) \rangle = \int f(\boldsymbol{\theta}') w(\boldsymbol{\theta} - \boldsymbol{\theta}') \, \mathrm{d}^2\theta' . \tag{4}$$

However, we note that this non-normalized estimator is expected to be more noisy than the one defined in Eq. (1) because of sampling noise. For example, even in the case of a flat field $f(\boldsymbol{\theta}) = 1$ measured without errors (so that $\hat{f}_n = 1$) we expect to have a noisy map if we use Eq. (3). For this reason, whenever possible the estimator (1) should be used instead.

The paper is organized as follows. In Sect. 2 we derive some preliminary expressions for the mean value of the map of Eq. (1). These results are generalized to a variable number N of objects in Sect. 3. If the weight function $w(\boldsymbol{\theta})$ is allowed to vanish, then some peculiarities arises. This case is considered in detail in Sect. 4. Section 5 is dedicated to general properties for the average of $\tilde{f}$ and related functions. In Sect. 6 we take an alternative approach which can be used to obtain an analytic expansion for $\tilde{f}$. In Sect. 7, we consider three specific examples of weight functions often used in practice. Finally, a summary of the results obtained in this paper is given in Sect. 8. A variation of the smoothing technique considered in this paper is briefly discussed in Appendix A.

2. Definitions and first results

Suppose one wants to measure an unknown field $f(\boldsymbol{\theta})$, a function of the "position" $\boldsymbol{\theta}$. [What $\boldsymbol{\theta}$ really means is totally irrelevant for our discussion. For example, $\boldsymbol{\theta}$ could represent the position of an object on the sky, the time of some observation, or the wavelength of a spectral feature. In the following, to focus on a specific case, we will assume that $\boldsymbol{\theta}$ represents a position on the sky and thus we will consider it as two-dimensional variable.] Suppose also that we can obtain a total of N *unbiased* estimates $\hat{f}_n$ for f at some points $\{\boldsymbol{\theta}_n\}$, and that each point can freely span a field Ω of surface A (Ω represents the area of the survey, i.e. the area where data are available). The points $\{\boldsymbol{\theta}_n\}$, in other words, are taken to be *independent* random variables with uniform probability distribution and density $\rho = N/A$ inside the set Ω of their possible values. We can then define the

smooth map of Eq. (1), or more explicitly

$$\tilde{f}(\boldsymbol{\theta}) = \frac{\sum_{n=1}^{N} \hat{f}_n w(\boldsymbol{\theta} - \boldsymbol{\theta}_n)}{\sum_{n=1}^{N} w(\boldsymbol{\theta} - \boldsymbol{\theta}_n)} . \tag{5}$$

In the rest of this paper we study the expectation value $\langle \tilde{f}(\boldsymbol{\theta}) \rangle$ of $\tilde{f}(\boldsymbol{\theta})$ (an alternative weighting scheme is briefly discussed in Appendix A). To simplify the notation we will assume, without loss of generality, that the weight function $w(\boldsymbol{\theta})$ is normalized, i.e.

$$\int_{\Omega} w(\boldsymbol{\theta}) \, \mathrm{d}^2\theta = 1 . \tag{6}$$

In order to obtain the ensemble average of $\tilde{f}$ we need to average over all possible measurements at each point, i.e. $\hat{f}_n$, and over all possible positions $\{\boldsymbol{\theta}_n\}$ for the N points. The first average is trivial, since $\tilde{f}(\boldsymbol{\theta})$ is linear on the data $\hat{f}_n$ and the data are unbiased, so that $\langle \hat{f}_n \rangle = f(\boldsymbol{\theta}_n)$. We then have

$$\langle \tilde{f}(\boldsymbol{\theta}) \rangle = \frac{1}{A^N} \int_{\Omega} \mathrm{d}^2\theta_1 \int_{\Omega} \mathrm{d}^2\theta_2 \cdots$$
$$\times \int_{\Omega} \mathrm{d}^2\theta_N \frac{\sum_{n=1}^{N} f(\boldsymbol{\theta}_n) w(\boldsymbol{\theta} - \boldsymbol{\theta}_n)}{\sum_{n=1}^{N} w(\boldsymbol{\theta} - \boldsymbol{\theta}_n)} . \tag{7}$$

Relabeling the integration variables we can rewrite this expression as

$$\langle \tilde{f}(\boldsymbol{\theta}) \rangle = \frac{N}{A^N} \int_{\Omega} \mathrm{d}^2\theta_1 \int_{\Omega} \mathrm{d}^2\theta_2 \cdots$$
$$\times \int_{\Omega} \mathrm{d}^2\theta_N \frac{f(\boldsymbol{\theta}_1) w(\boldsymbol{\theta} - \boldsymbol{\theta}_1)}{\sum_{n=1}^{N} w(\boldsymbol{\theta} - \boldsymbol{\theta}_n)} . \tag{8}$$

We now define a new random variable

$$y(\boldsymbol{\theta}) = \sum_{n=2}^{N} w(\boldsymbol{\theta} - \boldsymbol{\theta}_n) . \tag{9}$$

Note that the sum runs from $n = 2$ to $n = N$. Let us call $p_y(y)$ the probability distribution for $y(\boldsymbol{\theta})$. If we suppose that $\boldsymbol{\theta}$ is not close to the boundary of Ω, so that the support of $w(\boldsymbol{\theta} - \boldsymbol{\theta}')$ (i.e. the set of points $\boldsymbol{\theta}'$ where $w(\boldsymbol{\theta} - \boldsymbol{\theta}') \neq 0$) is inside Ω, then the probability distribution for $y(\boldsymbol{\theta})$ does not depend on $\boldsymbol{\theta}$. We anticipate here that below we will take the limit of large surveys, so that Ω tends to the whole plane, and $A \to \infty$, $N \to \infty$, such that $\rho = N/A$ remains constant. Since, by definition, the weight function is assumed to be non-negative, $p_y(y)$ vanishes for $y < 0$. Analogously, we call $p_w(w)$ the probability distribution for the weight w. These two probability distributions can be calculated from the equations

$$p_w(w) = \frac{1}{A} \int_{\Omega} \delta\big(w - w(\boldsymbol{\theta})\big) \, \mathrm{d}^2\theta , \tag{10}$$

$$p_y(y) = \frac{1}{A^{N-1}} \int_{\Omega} \mathrm{d}^2\theta_2 \cdots \int_{\Omega} \mathrm{d}^2\theta_N \delta(y - w_2 - \cdots - w_N)$$
$$= \int_0^{\infty} \mathrm{d}w_2 p_w(w_2) \int_0^{\infty} \mathrm{d}w_3 p_w(w_3) \cdots$$
$$\times \int_0^{\infty} \mathrm{d}w_N p_w(w_N) \delta(y - w_2 - \cdots - w_N) , \tag{11}$$

where δ is Dirac's distribution and where we have called $w_n = w(\boldsymbol{\theta}_n)$. Note that Eqs. (10) and (11) hold only if the N points $\{\boldsymbol{\theta}_n\}$ are uniformly distributed on the area A with density ρ, and if there is no correlation (so that the probability distribution for each point is $p_{\boldsymbol{\theta}}(\boldsymbol{\theta}_n) = 1/A$). Moreover, we are assuming here that the probability distribution for $y(\boldsymbol{\theta})$ does not depend on $\boldsymbol{\theta}$. This is true only if a given configuration of points $\{\boldsymbol{\theta}_n\}$ has the same probability as the translated set $\{\boldsymbol{\theta}_n + \boldsymbol{\phi}\}$. This translation invariance, clearly, cannot hold exactly for finite fields Ω; on the other hand, again, as long as $\boldsymbol{\theta}$ is far from the boundary of the field, the probability distribution for $y(\boldsymbol{\theta})$ is basically independent of $\boldsymbol{\theta}$. Note that in case of a field with masks, we also have to exclude in our analysis points close to the masks.

Using p_y we can rewrite Eq. (8) in a more compact form:

$$\langle \tilde{f}(\boldsymbol{\theta}) \rangle = \rho \int_{\Omega} \mathrm{d}^2 \theta_1 f(\boldsymbol{\theta}_1) w(\boldsymbol{\theta} - \boldsymbol{\theta}_1)$$
$$\times \int_0^{\infty} \frac{p_y(y)}{w(\boldsymbol{\theta} - \boldsymbol{\theta}_1) + y} \, \mathrm{d}y \, , \qquad (12)$$

where, we recall, $\rho = N/A$ is the density of objects. For the following calculations, it is useful to write this equation as

$$\langle \tilde{f}(\boldsymbol{\theta}) \rangle = \int_{\Omega} \mathrm{d}^2 \theta' f(\boldsymbol{\theta}') w(\boldsymbol{\theta} - \boldsymbol{\theta}') C\big(w(\boldsymbol{\theta} - \boldsymbol{\theta}')\big) \, \mathrm{d}^2 \theta' \, , \quad (13)$$

where $C(w)$ the *correcting factor*, defined as

$$C(w) = \rho \int_0^{\infty} \frac{p_y(y)}{w + y} \, \mathrm{d}y \, . \qquad (14)$$

Finally, we will often call the combination $w_{\mathrm{eff}}(\boldsymbol{\theta}) = w(\boldsymbol{\theta}) C\big(w(\boldsymbol{\theta})\big)$, which enters Eq. (13), *effective weight*.

Interestingly, Eq. (13) shows that the relationship between $\langle \tilde{f}(\boldsymbol{\theta}) \rangle$ and $f(\boldsymbol{\theta})$ is a simple convolution with the kernel w_{eff}. From the definition (5), we can also see that this kernel is normalized, in the sense that

$$\int_{\Omega} \langle \tilde{f}(\boldsymbol{\theta}) \rangle \, \mathrm{d}^2 \theta = \int_{\Omega} f(\boldsymbol{\theta}) \, \mathrm{d}^2 \theta \, . \qquad (15)$$

In fact, if we consider a "flat" signal, for instance $f(\boldsymbol{\theta}) = 1$, we clearly obtain $\langle \tilde{f}(\boldsymbol{\theta}) \rangle = 1$. On the other hand, from the properties of convolutions, we know that the ratio between the l.h.s. and the r.h.s. of Eq. (15) is constant, independent of the function $f(\boldsymbol{\theta})$. We thus deduce that this ratio is 1, i.e. that Eq. (15) holds in general. The normalization of w_{eff} will be also proved below in Sect. 5.1 using analytical techniques.

If $p_y(y)$ is available, Eq. (12) can be used to obtain the expectation value for the smoothed map $\tilde{f}$. In order to obtain an expression for p_y we use Markov's method (see, e.g., Chandrasekhar 1943; see also Deguchi & Watson 1987 for an application to microlensing studies). Let us define the Laplace transforms of p_y and p_w:

$$W(s) = \mathcal{L}[p_w](s) = \int_0^{\infty} \mathrm{e}^{-sw} p_w(w) \, \mathrm{d}w$$

$$= \frac{1}{A} \int_{\Omega} \mathrm{e}^{-sw(\boldsymbol{\theta})} \, \mathrm{d}^2 \theta \, , \qquad (16)$$

$$Y(s) = \mathcal{L}[p_y](s) = \int_0^{\infty} \mathrm{e}^{-sy} p_y(y) \, \mathrm{d}y = \big[W(s)\big]^{N-1} \, . \quad (17)$$

Hence p_y can in principle be obtained from the following scheme. First, we evaluate $W(s)$ using Eq. (16), then we calculate $Y(s)$ from Eq. (17), and finally we back-transform this function to obtain $p_y(y)$.

3. Continuous limit

So far we have considered a finite set Ω and a fixed number of objects N. In reality, one often deals with a non-constant number of objects, so that N is itself a random variable. [For example, if the objects we are studying are galaxies and if Ω is a small field, the expected number of galaxies in our field will follow a Poissonian distribution with mean value ρA, where ρ is the density of detectable galaxies in the sky.] Clearly, when we observe a field Ω we will obtain a particular value for the number of objects N inside the field. However, in order to obtain more general results, it is convenient to consider an ensemble average, and take the number of observed objects as a random variable; the results obtained, thus, will be averaged over all possible values of N.

A way to include the effect of a variable N in our framework is to note that, although we are observing a small area on the sky, each object could in principle be located at any position of the whole sky. Hence, instead of taking N as a random variable and A fixed, we consider larger and larger areas of the sky and take the limit $A \to \infty$. In doing this, we keep the object density $\rho = N/A$ constant. It is easily verified that the two methods (namely A fixed and N Poissonian random variable with mean ρA, or rather $A \to \infty$ with $\rho = A/N$ fixed), lead to the same results. In the following, however, we will take the latter scheme, and let A goes to infinity; correspondingly we take Ω as the whole plane.

From Eq. (16) we see that $W(s)$ is proportional to $1/A$. For this reason it is convenient to define the function

$$Q(s) = \int_{\Omega} \left[\mathrm{e}^{-sw(\boldsymbol{\theta})} - 1 \right] \mathrm{d}^2 \theta \, . \qquad (18)$$

If A is finite, we have $Q(s) = AW(s) - A$. Thus we can write, in the limit $A \to \infty$,

$$Y(s) = \lim_{N \to \infty} \left[1 + \frac{Q(s)\rho}{N} \right]^{N-1} = \mathrm{e}^{\rho Q(s)} \, . \qquad (19)$$

This equation replaces Eq. (17) when $N \to \infty$.

In order to further simplify the expression for the correcting factor $C(w)$, we rewrite its definition as

$$C(w) = \rho \int_0^{\infty} \frac{\zeta_w(x)}{x} \, \mathrm{d}x \, , \qquad (20)$$

where $x = y + w$, and the function ζ_w is defined as

$$\zeta_w(x) = \mathrm{H}(x - w) p_y(x - w) \, . \qquad (21)$$

Here $\mathrm{H}(x - w)$ is the Heaviside function at the position w, i.e.

$$\mathrm{H}(x) = \begin{cases} 0 & \text{if } x < 0 \, , \\ 1 & \text{otherwise.} \end{cases} \qquad (22)$$

The Laplace transform of ζ_w can be written as

$$Z_w(s) = \mathcal{L}[\zeta_w](s) = \mathrm{e}^{-ws}Y(s) . \tag{23}$$

In reality, for $C(w)$ we need to evaluate an integral over $\zeta_w(x)/x$. From the properties of the Laplace transform we have

$$\mathcal{L}\big[\zeta_w(x)/x\big](s) = \int_s^\infty Z_w(s')\,\mathrm{d}s' , \tag{24}$$

and thus we find

$$C(w) = \rho\mathcal{L}\big[\zeta_w(x)/x\big](0) = \rho\int_0^\infty Z_w(s')\,\mathrm{d}s'$$
$$= \rho\int_0^\infty \mathrm{e}^{-ws'}Y(s')\,\mathrm{d}s' = \mathcal{L}[\rho Y](w) . \tag{25}$$

This important result, together with Eqs. (18) and (19), can be used to readily evaluate the correcting factor. We note that, for our purposes, there is no need to evaluate the probability distribution p_y any more. This prevents us from calculating any inverse Laplace transform.

4. Vanishing weights

Since $w(\boldsymbol{\theta}) \geq 0$ for every $\boldsymbol{\theta} \in \Omega$, y is non-negative, so that $p_y(y) = 0$ if $y < 0$. In principle, however, we cannot exclude that the case $y = 0$ has a finite probability. In particular, for finite support weight functions, $p_y(y)$ could include the contribution from a Dirac delta distribution centered on zero.

Since y is the sum of weights at different positions, $y = 0$ can have a finite probability only if $w(\boldsymbol{\theta})$ vanishes at some points. In turn, if $P_0 = P(y = 0)$ is finite we could encounter situations where the numerator and the denominator of Eq. (5) vanish. In such cases we could not even define our estimator $\tilde{f}$. We also note that, in the continuous limit, P_0 is also the probability of having vanishing denominator in Eq. (5). As a result, if $P_0 > 0$, we have a finite probability of being unable to evaluate our estimator!

So far we have implicitly assumed that P_0 vanishes. Actually, we can explicitly evaluate this probability using the expression

$$P(y = 0) = \lim_{y \to 0} \int_0^y p_y(y')\,\mathrm{d}y' . \tag{26}$$

From the properties of the Laplace transform this expression can be written as

$$P(y = 0) = \lim_{s \to \infty} Y(s) = \lim_{s \to \infty} \mathrm{e}^{\rho Q(s)} . \tag{27}$$

We now observe that

$$\lim_{s \to \infty} \big[\mathrm{e}^{-sw(x)} - 1\big] = \begin{cases} -1 & \text{if } w \neq 0 , \\ 0 & \text{otherwise.} \end{cases} \tag{28}$$

As a result, in the limit considered, $-Q(s)$ approaches the area of the support of w. Calling this area π_w, we find finally

$$P(y = 0) = \mathrm{e}^{-\rho\pi_w} . \tag{29}$$

In the case where $w(\boldsymbol{\theta})$ has infinite support (for example if w is a Gaussian), P_0 vanishes as expected.

The result expressed by Eq. (29) can actually be derived using a more direct approach. Since $w(\boldsymbol{\theta}) \geq 0$, the condition $y = 0$ requires that all weights except $w(\boldsymbol{\theta}_1)$ are vanishing. In turn, this happens only if there is no object inside a neighborhood of the point considered. Since the area of this neighborhood is π_w, the number of objects inside this set follows a Poisson distribution of mean $\rho\pi_w$, and thus the probability of finding no object is given precisely by Eq. (29).

As mentioned before, $P_0 > 0$ is a warning that is same cases we cannot evaluate our estimator $\tilde{f}$. In order to proceed in our analysis allowing for finite-support weight functions, we decide to *explicitly exclude in our calculations cases where $w + y = 0$*: In such cases, in fact, we could not define $\tilde{f}$. In practice when smoothing the data we would mark as "bad points" the locations $\boldsymbol{\theta}$ where $\tilde{f}(\boldsymbol{\theta})$ is not defined. In taking the ensemble average, then, we would exclude, for each possible configuration $\{\boldsymbol{\theta}_n\}$, the bad points. In order to apply this prescription we need to modify p_y, the probability distribution for y, and explicitly exclude cases with $w + y = 0$. In other words, we define a new probability distribution for y given by

$$\tilde{p}_y(y) = \begin{cases} p_y(y) & \text{if } w \neq 0 , \\ [p_y(y) - P_0\delta(y)]/(1 - P_0) & \text{if } w = 0 . \end{cases} \tag{30}$$

Hence, if $w = 0$, we set to zero $P(y = 0)$ and then renormalize the distribution. An important consequence of the new prescription is that *the probability distribution for y no longer is independent of w*. In fact, the probability $P(y = 0)$ vanishes for $w = 0$, while is finite (for a finite-support weight) if $w \neq 0$. Using Eq. (30) in the definition of $Y(s)$ we then find (see Eqs. (17) and (19))

$$Y(s) = \mathrm{e}^{\rho Q(s)} - \Delta(w)\frac{P_0}{1 - P_0}\big[1 - \mathrm{e}^{\rho Q(s)}\big] , \tag{31}$$

where $\Delta(w)$ is 1 for $w = 0$ and vanishes for $w > 0$. Equation (31) replaces Eq. (19) for the cases when $P_0 > 0$, i.e. for weight functions with finite support.

In order to implement the new requirement $w + y \neq 0$, we still need to modify Eq. (12) [or, equivalently, Eq. (14)]. In fact, in deriving that result, we have assumed that all objects can populate the area Ω with uniform probability and in fact we have used a factor $1/A^N$ to normalize Eq. (7). Now, however, we must take into account the fact that objects cannot make a "void" around the point $\boldsymbol{\theta}$. As a result, we need a further factor $1/(1 - P_0)$ in front of Eqs. (12) and (14). In summary, the new set of equations is

$$Q(s) = \int_\Omega \big[\mathrm{e}^{-sw(\boldsymbol{\theta})} - 1\big]\,\mathrm{d}^2\theta , \tag{32}$$

$$Y(s) = \mathrm{e}^{\rho Q(s)} - \Delta(w)\frac{P_0}{1 - P_0}\big[1 - \mathrm{e}^{\rho Q(s)}\big] , \tag{33}$$

$$C(w) = \frac{\rho}{1 - P_0}\int_0^\infty \mathrm{e}^{-ws}Y(s)\,\mathrm{d}s , \tag{34}$$

where, we recall, $P_0 = \mathrm{e}^{-\rho\pi_w}$ vanishes for weight functions with infinite support π_w.

5. Properties

In this section we show a number of interesting properties of $C(w)$ and $\langle \tilde{f}(\boldsymbol{\theta}) \rangle$.

5.1. Normalization

As already shown [see above Eq. (13)], $\langle \tilde{f}(\boldsymbol{\theta}) \rangle$ is a simple convolution of $f(\boldsymbol{\theta})$ with w_{eff}. The normalization I of this convolution can be obtained from the expression

$$I = \int_\Omega w_{\text{eff}}(\boldsymbol{\theta}) \, \mathrm{d}^2\theta = \int_\Omega w(\boldsymbol{\theta}) C\bigl(w(\boldsymbol{\theta})\bigr) \, \mathrm{d}^2\theta \; . \tag{35}$$

From Eq. (34) we find

$$I = \frac{\rho}{1 - P_0} \int_\Omega \mathrm{d}^2\theta \int_0^\infty w(\boldsymbol{\theta}) \mathrm{e}^{-sw(\boldsymbol{\theta})} Y(s) \, \mathrm{d}s \; . \tag{36}$$

We first note that, in the general case, $Y(s)$ *depends* on $w(\boldsymbol{\theta})$. On the other hand, since the second term of $Y(s)$ is proportional to $\Delta(w)$ [see Eq. (33)], it does not contribute to the integral. We can then evaluate first the integral over $\boldsymbol{\theta}$, obtaining

$$\int_\Omega w(\boldsymbol{\theta}) \mathrm{e}^{-sw(\boldsymbol{\theta})} \, \mathrm{d}^2\theta = -Q'(s) \; , \tag{37}$$

and thus we have

$$I = -\frac{\rho}{1 - P_0} \int_0^\infty Q'(s) \mathrm{e}^{\rho Q(s)} \, \mathrm{d}s = \left. \frac{1}{1 - P_0} \mathrm{e}^{\rho Q(s)} \right|_0^\infty = 1 \; . \tag{38}$$

Hence we conclude that the estimator (5) is correctly normalized.

5.2. Scaling

It is easily verified that a simple scaling property holds for the effective weight. Suppose that we evaluate $w_{\text{eff}}(\boldsymbol{\theta})$ for a given weight function $w(\boldsymbol{\theta})$ and density ρ. If we rescale the weight function into $k^2 w(k\boldsymbol{\theta})$, and the density into $k^2\rho$, where k is a positive factor, the corresponding effective weight is also rescaled similarly to w, i.e. $k^2 w_{\text{eff}}(k\boldsymbol{\theta})$.

We also recall here that, although a normalized weight function has been assumed [see Eq. (6)], all results are clearly independent of the normalization of w. Hence, we can also consider a trivial scaling property: The weight function $k w(\boldsymbol{\theta})$ has effective weight $w_{\text{eff}}(\boldsymbol{\theta})$ independent of k.

We anticipate that the effective weight is very close to the original weight w for large densities ρ (see below Sect. 5.4). The scaling properties discussed above suggest that the shape of the effective weight is actually controlled by the expected number of objects for which the weight is significantly different from zero. This justifies the definition of the *weight area* $\mathcal{A}$ of w:

$$\mathcal{A} = \left[\int_\Omega w(\boldsymbol{\theta}) \, \mathrm{d}^2\theta \right]^2 \bigg/ \left[\int_\Omega \bigl[w(\boldsymbol{\theta}) \bigr]^2 \mathrm{d}^2\theta \right] \; . \tag{39}$$

The first factor in this definition ensures that $\mathcal{A}$ does not depend on the normalization of $w(\boldsymbol{\theta})$. It is easily verified that
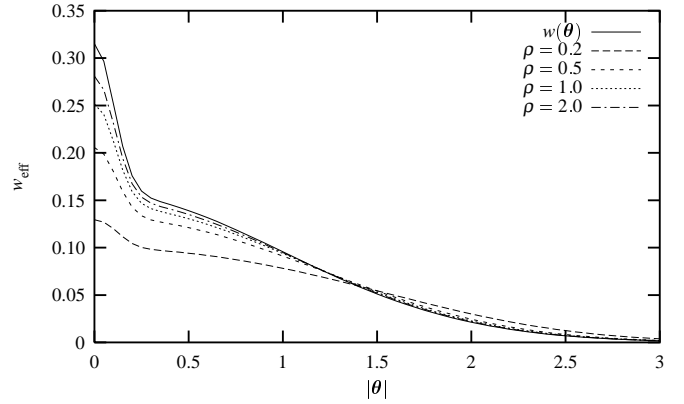
Fig. 1. The effective weight w_{eff} never exceed ρ. The original weight $w(\boldsymbol{\theta})$ is the combination of two Gaussians with different widths ($\sigma_1 = 1$ and $\sigma_2 = 0.1$). The central peak is severely depressed for relative low-densities. Note that the weight area is $\mathcal{A} \simeq 12.2$, so that a density of $\rho = 0.2$ corresponds to $\mathcal{N} \simeq 2.4$.

$\mathcal{A} = \pi_w$ for a top-hat weight function. Correspondingly, we define the *weight number of objects* as $\mathcal{N} = \rho\mathcal{A}$. Clearly, this quantity is left unchanged by the scalings considered above. Similar definitions can be provided for the effective weight w_{eff}. Explicitely, we have

$$\mathcal{A}_{\text{eff}} = \left[\int_\Omega w_{\text{eff}}(\boldsymbol{\theta}) \, \mathrm{d}^2\theta \right]^2 \bigg/ \left[\int_\Omega \bigl[w_{\text{eff}}(\boldsymbol{\theta}) \bigr]^2 \mathrm{d}^2\theta \right] \; , \tag{40}$$

and $\mathcal{N}_{\text{eff}} = \rho\mathcal{A}_{\text{eff}}$.

Numerical calculations for w_{eff} show clearly that $\mathcal{N}$, rather than ρ, is the key factor that controls how close the effective weight $w_{\text{eff}}(\boldsymbol{\theta})$ is to $w(\boldsymbol{\theta})$ (cf. Figs. 1, 3, and 4).

5.3. Behavior of $wC(w)$

We can easily study the general behavior of w_{eff}. We first consider the case of infinite-support weights ($P_0 = 0$).

We first note that $Y(s) > 0$ for every s, and thus $C(w)$ decreases as w increases. On the other hand, we have

$$w_{\text{eff}} = wC(w) = \rho Y(0) + \rho\mathcal{L}[Y'](w) \; . \tag{41}$$

Since $Y'(s) = \rho Q'(s) Y(s) < 0$ is negative, we have that $wC(w)$ increases with w. This result shows that the effective weight w_{eff} follows the general shape of the weight w, as expected. Moreover, the fact that $C(w)$ is monotonically decreasing implies that the effective weight w_{eff} is "broader" than w. For instance, we can say that there is a value w_1 for the weight such that $w_{\text{eff}} = wC(w)$ is not larger than w for $w > w_1$, and not smaller than w for $w < w_1$. In fact, since $C(w)$ is monotonic, the equation $C(w) = 1$ can have at most one solution. On the other hand, if $C(w) < 1$ (respectively, if $C(w) > 1$) for all w, then $w_{\text{eff}}(\boldsymbol{\theta}) < w(\boldsymbol{\theta})$ ($w_{\text{eff}}(\boldsymbol{\theta}) > w(\boldsymbol{\theta})$) for all $\boldsymbol{\theta}$. This inequality, however, cannot be true since both w and w_{eff} are normalized.

The property just shown can be used to derive a relation between the weight area $\mathcal{A}$ and the effective weight area $\mathcal{A}_{\text{eff}}$. Let us evaluate the integral

$$D = \int_\Omega \big[w(\boldsymbol{\theta}) + w_{\text{eff}}(\boldsymbol{\theta}) - 2w_1\big]\big[w(\boldsymbol{\theta}) - w_{\text{eff}}(\boldsymbol{\theta})\big] \, \mathrm{d}^2\theta \ . \quad (42)$$

This quantity is positive, since the integrand is positive (this is easily verified by distinguish the two cases $w(\boldsymbol{\theta}) > w_1$ and $w(\boldsymbol{\theta}) < w_1$, and by noting that both factors in the integrand have the same sign). On the other hand, if we expand the product we obtain

$$0 < D = \int_\Omega \big[w(\boldsymbol{\theta})\big]^2 \, \mathrm{d}^2\theta - \int_\Omega \big[w_{\text{eff}}(\boldsymbol{\theta})\big]^2 \, \mathrm{d}^2\theta$$
$$- 2w_1 \int_\Omega \big[w(\boldsymbol{\theta}) - w_{\text{eff}}(\boldsymbol{\theta})\big] \, \mathrm{d}^2\theta$$
$$= \mathcal{A}^{-1} - \mathcal{A}_{\text{eff}}^{-1} \ . \quad (43)$$

The last relation holds because of the normalization of $w(\boldsymbol{\theta})$ and $w_{\text{eff}}(\boldsymbol{\theta})$. We thus have shown that the effective weight area $\mathcal{A}_{\text{eff}}$ is always larger than the original weight area $\mathcal{A}$; analogously we have $\mathcal{N}_{\text{eff}} > \mathcal{N}$. This, clearly, is another indication that w_{eff} is "broader" than w.

It is also interesting to evaluate the limits of $wC(w)$ for small and large values of w. We have

$$\lim_{w\to\infty} wC(w) = \rho Y(0) = \rho \ . \quad (44)$$

Since we know that $wC(w)$ is monotonic, ρ is also a superior limit for the effective weight function. In other words, even if w has high peaks, the effective weight w_{eff} (that, we recall, is normalized) will never exceed the value ρ (see Fig. 1). We stress here that, since w_{eff} is normalized, its maximum value [which, in virtue of Eq. (44) does not exceed ρ] is a significant parameter. For example, using the relation $w_{\text{eff}}(\boldsymbol{\theta}) < \rho$ in the definition of $\mathcal{A}_{\text{eff}}$ we find immediately

$$\mathcal{A}_{\text{eff}}^{-1} = \int_\Omega \big[w_{\text{eff}}(\boldsymbol{\theta})\big]^2 \, \mathrm{d}^2\theta < \rho \int_\Omega w_{\text{eff}}(\boldsymbol{\theta}) \, \mathrm{d}^2\theta = \rho \ , \quad (45)$$

or, equivalently, $\mathcal{N}_{\text{eff}} > 1$. In other words, no matter how small $\mathcal{A}$ is, the effective weight will always "force" to use at least one object.

Equation (44) suggests also a *local* order-of-magnitude check for the effective weight: Assuming $w(\boldsymbol{\theta})$ normalized, we expect the effective weight to be significantly different from w for points where $w(\boldsymbol{\theta})$ is of the order of ρ or larger. Equation (45), instead, provides a *global* criterion: The effective weight will be significantly broader than w if $\mathcal{N}$ is of the order of unity or smaller. As anticipated above, thus, $\mathcal{N}$ is the real key factor that controls the shape of w_{eff} with respect to w.

The other limit for the effective weight is

$$\lim_{w\to 0^+} wC(w) = \rho \lim_{s\to\infty} Y(s) = 0 \ . \quad (46)$$

Thus, as expected, the effective weight w_{eff} vanishes as w vanishes.

If w has finite support, then the situation is slightly different. Given the definition of $Y(s)$, we expect for $C(w)$ a discontinuity for $w = 0$. Apart from this difference, the behavior of $C(w)$ and of $wC(w)$ is similar to the case considered above, namely $C(w)$ is monotonically decreasing and $wC(w)$ is increasing. We also have

$$\lim_{w\to\infty} wC(w) = \frac{\rho}{1 - P_0} \ , \quad (47)$$

which is similar to Eq. (44). Note that, as expected from the normalization of w_{eff}, $\pi_w \rho/(1 - P_0) \geq 1$ for all densities. We also find

$$\lim_{w\to 0^+} wC(w) = \frac{\rho P_0}{1 - P_0} \ . \quad (48)$$

In other words, the effective weight $wC(w)$ *does not vanish* for small w. If, instead, we take $w = 0$, then we have $wC(w) = 0$. This discontinuity, related to the term $\Delta(w)$ in Eq. (33), is a consequence of a number of properties for the effective weight: (i) w_{eff} is normalized; (ii) w_{eff} is broader than w; (iii) w_{eff} has the same support as w. We thus are forced to have a discontinuity for the effective weight.

The result obtained above is also convenient for simplifying equations for finite-support weight functions. In fact, for $w > 0$ we can clearly drop the last term of Eq. (33), thus recovering Eq. (19); if, instead, $w = 0$, we can directly set $wC(w) = 0$, without further calculations.

5.4. Limit of high and low densities

The final result considered in this section is the behavior of the correcting factor $C(w)$ in the limit $\rho \to \infty$ and $\rho \to 0$.

We observe that $Q(s) \leq 0$ and moreover $|Q(s)|$ increases with s. Thus, if s is large, $\mathrm{e}^{\rho Q(s)}$ vanishes quickly when $\rho \to \infty$. As a result, in the limit of large densities we are mainly interested in $Q(s)$ with s small. Expanding $Q(s)$ around $s = 0^+$, we find

$$Q(s) = \sum_{k=1}^\infty \frac{(-1)^k s^k S_k}{k!} \ , \quad (49)$$

where S_k is the k-th moment of w:

$$S_k = \int_\Omega \big[w(\boldsymbol{\theta})\big]^k \, \mathrm{d}^2\theta \ . \quad (50)$$

The normalization of w clearly implies $S_1 = 1$. Assuming $w > 0$, to first order we have then $Y(s) \simeq \mathrm{e}^{-s\rho}$, so that

$$C(w) \simeq \frac{1}{1 - P_0} \int_0^\infty \rho \mathrm{e}^{-sw} \mathrm{e}^{-s\rho} \, \mathrm{d}s = \frac{1}{1 - P_0} \frac{\rho}{\rho + w} \ . \quad (51)$$

This expression gives the correcting factor at the first order in $1/\rho$. If w has infinite support the expression reduces to $C(w) = 1/\big[1 + w/\rho\big]$. Note that the quantity $w/\rho \sim 1/\mathcal{N}$ is related the *weight number*, i.e. the expected number of objects which contribute significantly to the signal, for which the weight is not exceedingly small. Finally, at the zero order in the limit $\rho \to \infty$, $C(w)$ converges to unity. In this case the map $\tilde{f}(\boldsymbol{\theta})$ is expected to be a smoothing of $f(\boldsymbol{\theta})$ with the same weight function w.

In the limit $\rho \to 0$, we have $Y(s) \to 1$, and thus

$$C(w) \simeq \frac{\rho}{1 - P_0} \frac{1}{w} \ . \quad (52)$$
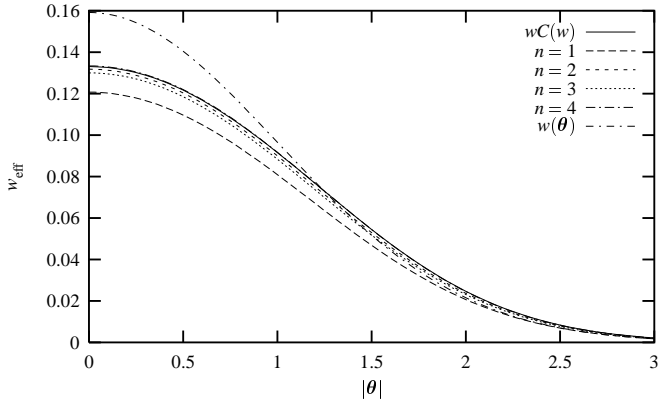
Fig. 2. The effective weight w_{eff} can be well approximated using the expansion (66). This plot shows the behavior of the n-th order expansion for a Gaussian weight function with unit variance (see Eq. (75)). The density used, $\rho = 0.5$, corresponds to $\mathcal{N} \simeq 6.3$. The convergence is already extremely good at the second order; for larger values of $\mathcal{N}$ the expansion converges more rapidly.

In the same limit, $P_0 \simeq 1 - \pi_w \rho$, so that we find

$$w_{\text{eff}} = wC(w) \simeq \frac{1}{\pi_w} . \tag{53}$$

Recalling that $wC(w)$ vanishes for $w = 0$, we conclude that the effective weight converges to a top-hat function with support π_w if $w(\boldsymbol{\theta})$ has finite-support (see Fig. 4 for an example).

6. Moments expansion

In the last section we have obtained an analytical expansion for $Q(s)$ which has then been used to obtain a first approximation for the correction function $C(w)$ valid for large densities ρ. Unfortunately, we have been able to obtain a simple result for $C(w)$ only to first order. Already at the second order, in fact, the correcting function would result in a rather complicated expression involving the error function erf.

Actually, there is a simpler approach to obtain an expansion of $C(w)$ at large ρ using the moments of the random variable y. Given the definition (9) for y, we expect that $\bar{y} \equiv \langle y \rangle$, the average value of this random variable, increases linearly with the density ρ of objects. Similarly, for large ρ the relative scatter $\langle (y - \bar{y})^2 \rangle / \bar{y}^2$ is expected to decrease. In fact, y is the sum of several independent random variables, and thus, in virtue of the central limit theorem, it must converge to a Gaussian random variable with appropriate average and variance.

Since the relative variance of y decreases with ρ, we can expand y in the denominator in Eq. (14), obtaining

$$C(w) = \rho \int_0^\infty \mathrm{d}y \, \frac{p_y(y)}{\bar{y} + w} \sum_{k=0}^\infty (-1)^k \left(\frac{y - \bar{y}}{\bar{y} + w} \right)^k$$

$$= \rho \sum_{k=0}^\infty (-1)^k \frac{1}{(\bar{y} + w)^{k+1}} \langle (y - \bar{y})^k \rangle , \tag{54}$$

where we have used the definition of the moments of y:

$$\bar{y} = \langle y \rangle = \int_0^\infty p_y(y) y \, \mathrm{d}y , \tag{55}$$

$$\langle (y - \bar{y})^k \rangle = \int_0^\infty p_y(y) (y - \bar{y})^k \, \mathrm{d}y . \tag{56}$$

In other words, if we are able to evaluate the moments of y we can obtain an expansion of $C(w)$. Actually, the "centered" moments can be calculated from the "un-centered" ones, defined by

$$\langle y^k \rangle = \int_0^\infty p_y(y) y^k \, \mathrm{d}y = (-1)^k Y^{(k)}(0) . \tag{57}$$

Here we have used the notation $Y^{(k)}(0)$ for the k-th derivative of $Y(s)$ evaluated at $s = 0$. Using Eq. (19) we can explicitly write the first few derivatives

$$Y(0) = 1 , \tag{58}$$

$$Y'(0) = \rho Q'(0) , \tag{59}$$

$$Y''(0) = \rho Q''(0) + \rho^2 [Q'(0)]^2 , \tag{60}$$

$$Y'''(0) = \rho Q'''(0) + 3\rho^2 Q''(0)Q'(0) + \rho^3 [Q'(0)]^3 , \tag{61}$$

$$Y^{(4)}(0) = \rho Q^{(4)}(0) + 4\rho^2 Q'''(0)Q'(0) + 3\rho^2 [Q''(0)]^2$$
$$+ 6\rho^3 Q''(0)[Q'(0)]^2 + \rho^4 [Q'(0)]^4 . \tag{62}$$

A nice point here is that, in principle, we can evaluate all the derivatives of $Y(s)$ in terms of derivatives of $Q(s)$ without any technical problem. Moreover, the derivatives of $Q(s)$ in zero are actually directly related to the moments of w. In fact we have

$$Q^{(k)}(0) = (-1)^k \int_\Omega [w(\boldsymbol{\theta})]^k \, \mathrm{d}^2\theta = (-1)^k S_k . \tag{63}$$

This simple relation allows us to express the moments of y in terms of the moments of w. For the first "centered" moments we find in particular

$$\bar{y} = \langle y \rangle = \rho , \qquad \langle (y - \bar{y})^2 \rangle = \rho S_2 , \tag{64}$$

$$\langle (y - \bar{y})^3 \rangle = \rho S_3 , \qquad \langle (y - \bar{y})^4 \rangle = \rho S_4 + 3\rho^2 S_2^2 . \tag{65}$$

Hence, we finally have

$$C(w) \simeq \frac{\rho}{\rho + w} + \frac{\rho^2 S_2}{(\rho + w)^3} - \frac{\rho^2 S_3}{(\rho + w)^4}$$
$$+ \frac{\rho^2 S_4 + 3\rho^3 S_2^2}{(\rho + w)^5} . \tag{66}$$

The first term if this expansion, $\rho/(\rho + w)$, has already been obtained in Eq. (51). Other terms represents higher order corrections to $C(w)$. In Fig. 2 we show the result of applying this expansion to a Gaussian weight.

In closing this section we note that, regardless to the value of π_w, $P_0 = \mathrm{e}^{-\rho\pi_w}$ is vanishing at all orders for $\rho \to \infty$, and thus we cannot see the peculiarities of finite-support weight functions here.

7. Examples

In this section we consider three typical examples of weight functions, namely a top-hat function, a Gaussian, and a parabolic weight function. For simplicity, in the following we will consider weight functions with fixed "width." The results obtained can then be adapted to weight functions with different widths using the scaling property (Sect. 5.2).

7.1. Top-hat

The simplest case we can consider for $w(\boldsymbol{\theta})$ is a top-hat function of unit radius, which can be written as

$$w(\boldsymbol{\theta}) = \frac{1}{\pi} \mathrm{H}(1 - |\boldsymbol{\theta}|) . \tag{67}$$

In this case we immediately find for $w > 0$

$$Q(s) = \pi(\mathrm{e}^{-s/\pi} - 1) , \tag{68}$$

$$Y(s) = \exp[\pi\rho(\mathrm{e}^{-s/\pi} - 1)] . \tag{69}$$

We now note that since $w(\boldsymbol{\theta})$ is either 0 or $1/\pi$, we just need to evaluate $C(1/\pi)$. We then find

$$C(1/\pi) = \frac{\rho}{1 - P_0} \int_0^\infty \mathrm{e}^{-s/\pi} Y(s)\, \mathrm{d}s = 1 , \tag{70}$$

as expected.

For the top-hat function we can also explicitly obtain the probability distribution for y. If $w > 0$ we have

$$p_y(y) = \sum_{n=0}^\infty \frac{\mathrm{e}^{-\rho\pi}(\rho\pi)^n}{\pi n!} \delta(y - n/\pi) , \tag{71}$$

and thus

$$
\begin{aligned}
C(w) &= \frac{\rho}{1 - P_0} \int_0^\infty \frac{p_y(y)}{w + y}\, \mathrm{d}y \\
&= \frac{\rho P_0}{\pi(1 - P_0)} \sum_{n=0}^\infty \frac{(\rho\pi)^n}{(w + n/\pi)n!} .
\end{aligned}
\tag{72}
$$

From this expression we easily obtain $C(1/\pi) = 1$. Moreover, we can evaluate the two limits

$$\lim_{w\to\infty} wC(w) = \frac{\rho}{1 - P_0} , \tag{73}$$

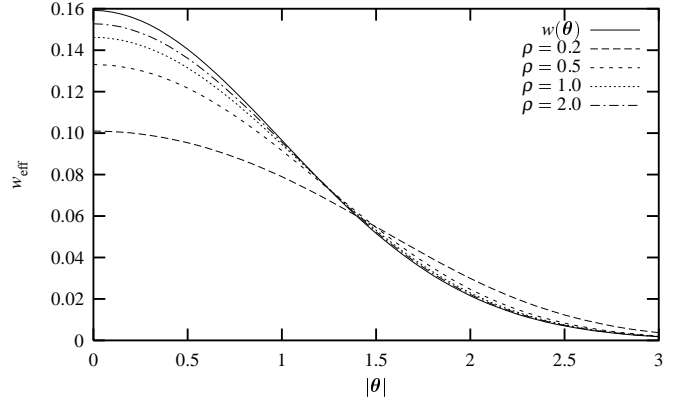$$\lim_{w\to 0^+} wC(w) = \frac{\rho P_0}{1 - P_0} , \tag{74}$$

thus regaining the results of Sect. 5.3.

7.2. Gaussian

A weight function commonly used is a Gaussian of the form

$$w(\boldsymbol{\theta}) = \frac{1}{2\pi} \exp(-|\boldsymbol{\theta}|^2/2) . \tag{75}$$

Unfortunately, we cannot explicitly integrate $Q(s)$ and thus we are unable to obtain a finite expression for $C(w)$. The results of a numerical calculations are however shown in Fig. 3.

Fig. 3. Effective weight function corresponding to a Gaussian weight function. The original weight function is a normalized Gaussian of unit variance, with weight area $\mathcal{A} \simeq 12.5$. Significantly broader effective weights are obtained if the weight number is smaller than $\mathcal{N} < 10$.
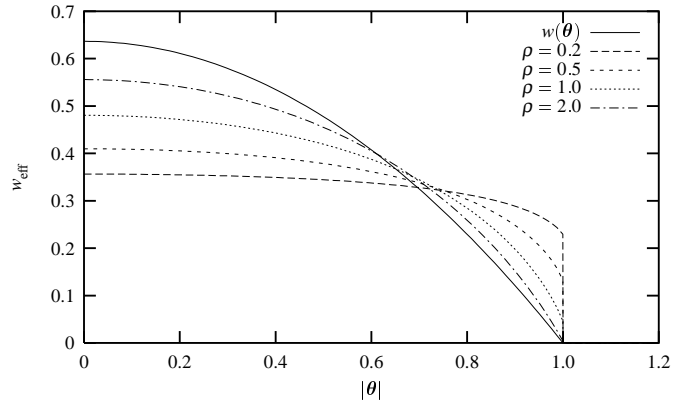
Fig. 4. Effective weight function corresponding to a Gaussian weight function. Note the discontinuity at $|\boldsymbol{\theta}| = 1$, corresponding to the boundary of the support of w. The original weight function is a normalized parabolic function with weight area $\mathcal{A} \simeq 2.4$. Note that this weight area is significantly smaller than the ones encountered in previous examples. For low densities, w_{eff} converges to a top-hat function, in accordance with the results of Sect. 5.4.

7.3. Parabolic weight

As the last, example we consider a parabolic weight function with expression

$$w(\boldsymbol{\theta}) = \begin{cases} \dfrac{2}{\pi}(1 - |\boldsymbol{\theta}|^2) & \text{for } |\boldsymbol{\theta}| < 1 , \\ 0 & \text{otherwise} . \end{cases} \tag{76}$$

We then find

$$Q(s) = \frac{1 - \mathrm{e}^{-2s/\pi}}{2s} - \pi . \tag{77}$$

Unfortunately, we cannot proceed analytically and determine $C(w)$. We thus report the results of numerical integrations in

Fig. 4. Note that, as expected, the resulting effective weight has a discontinuity at $|\boldsymbol{\theta}| = 1$.

8. Conclusions

In this paper we have studied from a statistical point of view the effect of smoothing irregularly sampled data. The main results obtained can be summarized in the following points.

1. The mean smooth map, $\langle \tilde{f}(\boldsymbol{\theta}) \rangle$, is a *convolution* of the unknown field $f(\boldsymbol{\theta})$ with an *effective weight* w_{eff}.
2. We have provided simple expressions to evaluate the effective weight. These expressions can be easily used, for example, to obtain numerical estimates of w_{eff}.
3. The effective weight $w_{\text{eff}}(\boldsymbol{\theta})$ and the weight function $w(\boldsymbol{\theta})$ share the same support and have a similar "shape." However, w_{eff} is broader than w, expecially for low densities of objects; moreover, w_{eff} has a discontinuity on the boundary of the support.
4. We have shown that the density of objects ρ (or $\rho/(1 - P_0)$ for finite-support weight functions) is a natural upper limit for the effective weights.
5. The weight number $\mathcal{N}$ has been shown to be the key factor that controls the convergence of w_{eff} to w. We have also shown that $\mathcal{N}_{\text{eff}} > 1$.
6. The effective weight converges to $w(\boldsymbol{\theta})$ for large densities ρ, and to a top-hat function for low densities.
7. We have provided an analytic expansion for w_{eff} which is shown to converge quickly to the exact weight function.
8. Finally, we have considered three typical examples and shown the behavior of w_{eff} for different densities.

Given the wide use in astronomy of the smoothing technique considered in this paper, an exact statistical characterization of the expectation value of the smoothed map is probably interesting *per se*.

Finally, we notice that other methods different from Eqs. (5) or (3) can be used to obtain continuous maps from irregularly sampled data. In particular, triangulation techniques can represents an interesting alternative to the simple weighted average considered here (see, e.g., Bernardeau & van de Weygaert 1996; Schaap & van de Weygaert 2000).

Appendix A: Object intrinsic weight

The simple weighting scheme considered in this paper [see Eq. (5)] is actually the base of several similar schemes with slightly different properties. In this paper, for simplicity and clarity, we have confined the discussion to the simple case, which is already rich of peculiarities and unexpected results (for example, the behavior in case of vanishing weights). In this appendix, however, we will describe a slightly more complicated estimator often used in astronomy.

Suppose, as for the Eq. (5), that we can measure a given field $f(\boldsymbol{\theta})$ at some positions $\boldsymbol{\theta}_n$ inside a set Ω. Suppose also that we decide to use, for each object, a weight $u_n > 0$ which

is independent of the position $\boldsymbol{\theta}_n$ (the weight u_n could however depend on other intrinsic properties of the object, such as its luminosity or angular size). We then replace Eq. (5) with

$$\tilde{f}(\boldsymbol{\theta}) = \frac{\sum_{n=1}^{N} \hat{f}_n u_n w(\boldsymbol{\theta} - \boldsymbol{\theta}_n)}{\sum_{n=1}^{N} u_n w(\boldsymbol{\theta} - \boldsymbol{\theta}_n)} \ . \tag{A.1}$$

A typical situation for which this estimator is useful, is when some error estimates $\{\sigma_n\}$ are available for each object. Then, if we set $u_n = 1/\sigma_n^2$, we obtain an estimator which optimizes the signal-to-noise ratio.

Since the quantities $\{u_n\}$ do not depend on the position, we can study the statistical properties of the estimator (A.1) provided that the probability distribution $p_u(u)$ of u is available. In particular we obtain [cf. Eq. (8)]

$$\langle \tilde{f}(\boldsymbol{\theta}) \rangle = \frac{N}{A^N} \int_\Omega \mathrm{d}^2\theta_1 \int_0^\infty \mathrm{d}u_1 \, p_u(u_1) \cdots$$
$$\times \int_\Omega \mathrm{d}^2\theta_N \int_0^\infty \mathrm{d}u_N \, p_u(u_N) \frac{f(\boldsymbol{\theta}_1) u_1 w(\boldsymbol{\theta} - \boldsymbol{\theta}_1)}{\sum_{n=1}^{N} u_n w(\boldsymbol{\theta} - \boldsymbol{\theta}_n)} \ . \tag{A.2}$$

Similarly to Eq. (9), we define

$$y(\boldsymbol{\theta}) = \sum_{n=2}^{N} u_n w(\boldsymbol{\theta} - \boldsymbol{\theta}_n) = \sum_{n=2}^{N} v_n \ , \tag{A.3}$$

with $v_n = u_n w(\boldsymbol{\theta}_n)$. Using y and $\{v_n\}$ and defining the probability distributions $p_y(y)$ and $p_v(v)$, we can write the analogous of Eqs. (10) and (11). In this way we obtain results similar to Eqs. (12-15) that we do not report here. Finally then we have

$$V(s) = \mathcal{L}[p_v](s) = \int_0^\infty \mathrm{e}^{-sv} p_v(v) \, \mathrm{d}v$$
$$= \frac{1}{A} \int_0^\infty p_u(u) \, \mathrm{d}u \int_\Omega \mathrm{e}^{-suw(\boldsymbol{\theta})} \, \mathrm{d}^2\theta \ , \tag{A.4}$$

$$Y(s) = \mathcal{L}[p_y](s) = \int_0^\infty \mathrm{e}^{-sy} p_y(y) \, \mathrm{d}y = [V(s)]^{N-1} \ . \tag{A.5}$$

In other words, we basically recover Eqs. (16) and (17) with the significant difference that now the Laplace transform $V(s)$ of $p_v(v)$ is given by the superpositions of functions like $W(us)$ weighted with the probability distribution $p_u(u)$. Note that all functions $W(us)$ have the same shape but differ for the scaling of the independent variable. In case where $p_u(u) = \delta(u - 1)$ is Dirac's delta distribution centered in 1 we exactly reproduce Eqs. (16) and (17), as expected.

The continuous limit (cf. Sect. 3) does not present particular difficulties. The first significant change regards Eq. (18), which now becomes

$$R(s) = \int_0^\infty p_u(u) \, \mathrm{d}u \int_\Omega \left[\mathrm{e}^{-suw(\boldsymbol{\theta})} - 1 \right] \mathrm{d}^2\theta$$
$$= \int_0^\infty p_u(u) Q(us) \, \mathrm{d}u \ . \tag{A.6}$$

Correspondingly, $Y(s) = \exp[\rho R(s)]$. Finally, writing the effective weight as $w_{\text{eff}} = wC(w)$, we have

$$C(w) = \int_0^\infty p_u(u) B(uw) \, \mathrm{d}u \ , \tag{A.7}$$

$$B(w) = \mathcal{L}[\rho Y](w) \ . \tag{A.8}$$

The equations written here can be used to practically evaluate $B(w)$ (and thus $w_{\mathrm{eff}}(\boldsymbol{\theta})$), but also to derive the properties of the effective weight, as done in for the simple estimator (5). Here we do not carry out the calculations, since they are straightforward modifications of the calculations of Sect. 5; moreover the results obtained are basically identical to the ones reported in that section.

References

Bernardeau F., van de Weygaert R., 1996, MNRAS 279, 693
Chandrasekhar S., 1943, Rev. Mod. Phys. 15, 1
Deguchi S., Watson W.D., 1987, Phys. Rev. Let. 59, 2814
Dekel A., Bertschinger E., Faber S. M., 1990, ApJ 364, 349
Kolaczyk E. D., Dixon D. D., 2000, ApJ 534, 490
Lombardi M., Bertin G., 1998, A&A 335, 1
Schaap W. E., van de Weygaert R., 2000, A&A 363, L29
van Waerbeke L., 2000, MNRAS 313, 524